

ПОСЛЕДОВАТЕЛЬНАЯ ОЦЕНКА НАИМЕНЬШИХ КВАДРАТОВ ЛИНЕЙНЫХ ПАРАМЕТРОВ ARCH-ПРОЦЕССА

Для гарантированной оценки параметров устойчивого процесса авторегрессии с условной гетероскедастичностью ARCH(q) методом наименьших квадратов в данной работе предложено использовать специальное правило останова, которое существенно зависит от поведения минимального собственного значения наблюдаемой информационной матрицы по Фишеру. Определена асимптотика верхней границы для среднеквадратичного уклонения оценок и асимптотическая формула среднего значения момента останова наблюдений.

Рассмотрим устойчивую модель с условной гетероскедастичностью ARCH(q) на вероятностном пространстве (Ω, F, P)

$$x_k = \theta_1 x_{k-1} + \dots + \theta_q x_{k-q} + \sqrt{1 + \sigma_1^2 x_{k-1}^2 + \dots + \sigma_q^2 x_{k-q}^2} \varepsilon_k, \quad (1)$$

где $(\varepsilon_k)_{k \geq 1}$ – последовательность независимых одинаково распределенных (н.о.р.) стандартных гауссовских случайных величин. Вектор начальных значений $X_0 = (x_0, \dots, x_{-q+1})'$ случайный с $E_0 \|X_0\|^2 < +\infty$ и не зависит от последовательности $(\varepsilon_k)_{k \geq 1}$, кроме того, имеет непрерывное распределение $\mu(x)$. Здесь штрих означает транспонирование.

Для использования оценок по методу наименьших квадратов (МНК) удобно ввести статистику

$$\bar{L}_k = \sqrt{1 + \sigma_1^2 x_{k-1}^2 + \dots + \sigma_q^2 x_{k-q}^2}, \\ L_k = \sqrt{1 + x_{k-1}^2 + \dots + x_{k-q}^2}, k \in N$$

и поделить на первую обе части уравнения (1). Далее мы будем строить процедуру, основанную на следующей модификации оценки МНК параметра θ :

$$\theta(n) = M_n^{-1} \sum_{k=1}^n X_{k-1} x_k L_{k-1}^{-1}, \quad M_n = \sum_{k=1}^n X_{k-1} X_{k-1}', \quad (2)$$

где $X_k = L_{k-1}^{-1} (x_k, \dots, x_{k-q+1})'$, M_n^{-1} обозначает матрицу, обратную к матрице M_n , если $\det M_n > 0$ и $M_n^{-1} = 0$ в противном случае.

Процесс (1) и его обобщения (GARCH, TARARCH) широко используются при анализе временных рядов, в частности финансовых. Асимптотические свойства модификации оценки МНК (2) этого процесса сильно зависят от значений неизвестных параметров θ .

Чтобы осветить проблему гарантированного оценивания процесса авторегрессии с условной гетероскедастичностью, рассмотрим модель авторегрессии первого порядка (или ARCH(1) при $\sigma = 0$) $x_i = \theta x_{i-1} + \varepsilon_i$, $i \geq 1$.

В [2] предложена следующая последовательная оценка МНК:

$$\theta^*(h) = h^{-1} \left(\sum_{k=1}^{\tau(h)-1} x_{k-1} x_k + \bar{\beta}(h) x_{\tau(h)-1} x_{\tau(h)} \right), \quad (3)$$

где

$$\tau(h) = \inf \left\{ n \geq 1 : \sum_{i=1}^n x_{i-1}^2 \geq h \right\}, \quad h > 0 \quad (4)$$

и $\bar{\beta}(h)$ – корректирующий множитель, заданный уравнением

$$\sum_{k=1}^{\tau(h)-1} x_{k-1}^2 + \bar{\beta}(h) x_{\tau(h)-1}^2 = h \quad (5)$$

(это приводит к тому, что $0 < \bar{\beta}(h) \leq 1$). Показано, что $E_0 \theta^*(h) = \theta$, $\theta \in R$, т.е. $\theta^*(h)$ является несмещенной оценкой θ , и $\sup_{\theta \in R} E_0 (\theta^*(h) - \theta)^2 \leq \sigma^2 / h$, что означает, что $\theta^*(h)$ является гарантированной в среднеквадратичном смысле равномерно по $\theta \in (-\infty, \infty)$.

Последовательные схемы выборок, описанные выше, не могут быть применены при построении гарантированных оценок на основе МНК в случае AR(q), ARCH(q) порядка $q > 1$. Задачи гарантированного оценивания параметров в AR(q) и

более сложных моделей случайной регрессии решались в два этапа, что требовало несколько (случайное число) оценок наименьших квадратов (2) [2, 7].

В данной работе для модели (1) мы построим одноэтапную последовательную оценку наименьших квадратов. Основным результатом состоит в том, что при $h \rightarrow \infty$

$$\sup_{\theta \in K} E_0 \|\theta^*(h) - \theta\|^2 \leq \frac{a_K}{h} (1 + o(1)),$$

где K – компактное множество в особой области устойчивости процесса (1), которую определим позже; a_K – некоторая константа, известная, если заданы параметры дисперсии, в противном случае – константа a_K не определена.

МАРКОВСКИЙ ПОДХОД. РАВНОМЕРНАЯ ЭРГОДИЧНОСТЬ ПРОЦЕССА ARCH

Чтобы исследовать модель (1), необходимо ввести следующие случайные матрицы $q \times q$. Определим

$$A_n = \begin{pmatrix} \alpha_1(n) & \dots & \alpha_q(n) \\ I_{q-1} & & 0 \end{pmatrix},$$

где I_{q-1} – единичная матрица порядка $q-1$; $\alpha_i(n) = \theta_i + \sigma_i \eta_i(n)$, $1 \leq i \leq q$ ($\eta_i(n)$, $1 \leq i \leq q$, $n \geq 1$) – н.о.р. стандартные гауссовские.

Мы будем предполагать, что для (1) выполняется следующее условие **H**₁: собственные значения матрицы $E_{0,\sigma} A_1 \otimes A_1$ по модулю не превосходят единицы, где $\otimes$ означает Кронекерово произведение матриц.

Вектор линейных параметров $\theta = (\theta_1, \dots, \theta_q)'$ неизвестен, и θ_i , σ_i такие, что выполняется условие **H**₁. Пусть Λ_σ означает множество всех векторов $\theta = (\theta_1, \dots, \theta_q)'$ с такими координатами θ_i , что выполняется условие **H**₁ при фиксированном векторе коэффициентов дисперсии $\sigma = (\sigma_1, \dots, \sigma_q)'$.

Согласно [6], модель (1) можно переписать в форме авторегрессии со случайными коэффициентами

$$y_k = (\theta_1 + \sigma_1 \eta_1(k)) y_{k-1} + (\theta_q + \sigma_q \eta_q(k)) y_{k-q} + \xi_k, \quad k = 1, 2, \dots, \quad (6)$$

где $(\xi_k)_{k \geq 1}$, $(\eta_i(k))_{k \geq 1, 1 \leq i \leq q}$ – последовательности н.о.р. стандартных гауссовских случайных величин. Тогда процесс $(y_k)_{k \geq 1}$ будет одинаково распределен с процессом $(x_k)_{k \geq 1}$. В векторной форме уравнение (6) переписывается в виде

$$Y_k = A_k Y_{k-1} + \zeta_k, \quad k = 1, 2, \dots, \quad (7)$$

где $Y_k = (y_k, \dots, y_{k-q+1})'$, $\zeta_k = (\xi_k, 0, \dots, 0)'$, матрицы A_k определены выше. Выразим Y_k через Y_0 и матрицы A_k :

$$Y_k = A_k \dots A_1 Y_0 + \sum_{s=0}^{k-2} A_k \dots A_{k-s} \zeta_{k-s-1} + \zeta_k, \quad k = 1, 2, \dots \quad (8)$$

Условие **H**₁ гарантирует устойчивость процесса $(y_k)_{k \geq 1}$, а с ним и $(x_k)_{k \geq 1}$. Действительно, оно влечет геометрическую скорость сходимости норм произведения

матриц A_k (см. [1], [5]) в среднеквадратическом смысле, т.е. найдутся такие константы $c > 0$, $0 < \rho < 1$, что

$$\sup_{\theta \in K} E_\theta \|A_k \cdots A_1\|^2 \leq c^2 \rho^{2k}, \quad k = 1, 2, \dots, \quad (9)$$

где K содержится в Λ_σ . Это следует из непрерывности собственных значений $EA_1 \otimes A_1$ по θ . Тогда рассматриваемые нормы сходятся с геометрической скоростью почти всюду. Откуда следует, что

$$\lim_k \|Y_k\| < +\infty, \quad k = 1, 2, \dots, P_\theta - \text{п.н.} \quad (10)$$

Согласно [8], процесс Y_n является марковским.

Лемма 1. Марковская цепь (7) является ψ -неразложимой, где ψ – максимальная мера неразложимости. Кроме того, $\sup \psi = R^q$.

Доказательство. Утверждения леммы следуют из разложения (8), результата из [8] и определения максимальной меры неразложимости и носителя меры в [8].

Лемма 2. Марковская цепь (7) при условии $\mathbf{H}_1$ является аperiodической позитивной Харрисовой цепью.

Доказательство. Можно показать, что марковская цепь (7) представляет собой аperiodическую T -цепь (см [8]). Тогда утверждения леммы следуют из условия $\mathbf{H}_1$, (10), леммы 1 и результатов из [8].

В дальнейшем для построения гарантированной оценки линейных параметров потребуется следующее важное свойство марковской цепи (7).

Теорема 1. Процесс (7) при условии $\mathbf{H}_1$ является $1 + \|X\|^2$ -равномерно эргодическим, причем для него выполняется условие (V4) из [8] и величина $1 - \beta$ в этом условии равномерно по $\theta \in K$ ограничена сверху константой строго меньшей единицы, где K – компакт в области Λ_σ .

Доказательство. Теорему можно доказать, используя условия равномерной эргодичности из [8] и леммы 1, 2. Последнее утверждение теоремы следует из непрерывности собственных значений матрицы $E_{\theta, \sigma} A_1 \otimes A_1$ по θ .

ОЦЕНИВАНИЕ ЛИНЕЙНЫХ ПАРАМЕТРОВ МОДЕЛИ ARCH(q)

Цель этого раздела – показать, что разработанная схема оценивания, основанная на оценке наименьших квадратов, может быть использована для определения линейных параметров (1) с заданной точностью.

Утверждение 1. Для модели (1) выполнено условие $\lambda_q(M_n) \rightarrow \infty$, при $n \rightarrow \infty$ P_θ -п.н., где $\lambda_q(A)$ обозначает минимальное по модулю собственное значение матрицы A .

Доказательство. Утверждение можно доказать, используя непрерывную зависимость корней многочлена от его коэффициентов и положительную определенность матрицы F (см. [1]).

Введем случайную величину

$$n_0 = \inf \{l \geq 1 : \lambda_q(M_l) > 0\},$$

которая конечна P_θ -п.н. в силу утверждения 1. Последовательная оценка МНК линейных параметров процесса (1) определяется следующим образом:

$$\begin{aligned} \theta^*(h) &= \tilde{M}_{\tau(h)}^{-1} \sum_{k=1}^{\tau(h)} \beta_k X_{k-1} L_{k-1}^{-1} x_k; \\ \tilde{M}_{\tau(h)}^{-1} &= \sum_{k=1}^{\tau(h)} \beta_k X_{k-1} X_{k-1}' ; \end{aligned} \quad (11)$$

$$\tau(h) = \inf \left\{ n \geq n_0 : \|M_n^{-2}\|^{1/2} \leq h^{-1} \right\}, \quad (12)$$

M_n задается (2). Величина β_k находится из соотношения

$$\beta_k = \begin{cases} 1, & \text{если } k < \tau(h), \\ \bar{\beta}(h), & \text{если } k = \tau(h), \end{cases} \quad \text{а } \bar{\beta}(h) \text{ задана уравнением}$$

$$\left\| \sum_{k=1}^{\tau(h)-1} X_{k-1} X_{k-1}' + \bar{\beta}(h) X_{\tau(h)-1} X_{\tau(h)-1}' \right\|^{1/2} = h^{-1}.$$

Утверждение 2. P_θ -п.н. ($\theta \in K$) выполняется предельное соотношение $\lim_n \frac{M_n}{n} = F = E_\theta L_0^{-2} \tilde{Y}_0 \tilde{Y}_0'$, где F – не случайная, положительно определенная матрица, L_0 определяется на базе вектора $\tilde{Y}_0$, $\tilde{Y}_n$ – стационарная версия процесса Y_n .

Доказательство. Результат следует из леммы 2 (см. [8]), а положительная определенность F доказывается, как в [1].

Лемма 3. Стационарный процесс (7) удовлетворяет условию сильного перемешивания с геометрической скоростью сходимости, где соответствующий коэффициент пропорционален s^n и $s < 1$ равномерно по $\theta \in K$.

Доказательство. Утверждение леммы следует из леммы 1, теоремы 1 и результатов из [8].

Лемма 4. Введем функцию $G(X) = X(1 + XX')^{-1}$. Положим $Z_n = G(Y_n)$, $\tilde{Z}_n = G(\tilde{Y}_n)$. Тогда, найдется такая константа $C > 0$, что

$$\sup_{\theta \in K} E_\theta \|Z_k - \tilde{Z}_k\| \|Z_l - \tilde{Z}_l\| < C \rho^k \rho^l, \quad (13)$$

где K – компакт в области Λ_σ .

Доказательство. Лемму можно доказать, используя разложение в ряд Тейлора нормы разности между процессом (7) и его стационарной аппроксимацией, а также разложение (8) и неравенства (9).

$$\text{Лемма 5. } \sup_{\theta \in K} E_\theta \left(\sum_{k=0}^{\infty} \|Z_k Z_k' - \tilde{Z}_k \tilde{Z}_k'\| \right)^2 < +\infty.$$

Доказательство. Утверждение леммы получим, используя лемму 4 и теорему о мажорируемой сходимости.

Лемма 6. Для любого компакта $K \subset \Lambda_\sigma$

$$\sup_{\theta \in K} \|F^{-2}\|^{1/2} < +\infty$$

и функции от θ $\|F^{-2}\|^{1/2}$ равномерно непрерывны на K .

Доказательство. Используя результат из [6], неравенство для симметричных матриц $\|EM\|^2 \leq E\|M\|^2$, неравенства (9) и (13) можно показать, что F непрерывна по θ на K . $\|F^{-2}\|^{1/2}$ выражается через собственные числа F , которые также непрерывны по θ на K .

Лемма 7. Для всех $\eta > 0$ имеет место соотношение

$$\sum_{n=1}^{\infty} \sup_{\theta \in K} P_\theta \left(\left\| \frac{1}{n} \sum_{k=0}^n (\tilde{Z}_k \tilde{Z}_k' - F) \right\| > \eta \right) < +\infty.$$

Доказательство. Используя неравенство Коши–Буняковского, леммы 3, 5, неравенство для ковариаций и коэффициентов сильного перемешивания (см., например, [4]) и теорему о больших отклонениях [3]), можно получить утверждение леммы.

Лемма 8. Для всех $h > 0$

$$E_{\theta} \operatorname{tr} \tilde{M}_{\tau} \leq E_{\theta} \tau. \quad (14)$$

Доказательство. Доказательство сразу получается, если расписать величину $\operatorname{tr} \tilde{M}_{\tau}$.

Основные свойства гарантированной оценки приведены в следующих теоремах.

Теорема 2. Пусть процесс (1) устойчивый, и его линейные коэффициенты лежат в области Λ_{σ} . Тогда для всех компактных множеств $K \subset \Lambda_{\sigma}$

$$\limsup_{h \rightarrow \infty} \sup_{\theta \in K} \left| E_{\theta} \frac{\tau(h)}{h} - \|F^{-2}\|^{1/2} \right| = 0.$$

Доказательство. Сначала с помощью лемм 6 и 7 можно показать, что $\limsup_{h \rightarrow \infty} \sup_{\theta \in K} E_{\theta} \frac{\tau(h)}{h} < +\infty$. Для этого

достаточно сходимости ряда $\sum_{n=1}^{\infty} \sup_{\theta \in K} P_{\theta} \left(\left\| \frac{M_n}{n} - F \right\| > \eta \right)$.

Дальнейшее доказательство проводится с использованием леммы 7.

Теорема 3. При условиях теоремы 2 для всех компактных множеств $K \subset \Lambda_{\sigma}$

$$\sup_{\theta \in K} E_{\theta} \left\| \theta^*(h) - \theta \right\|^2 \leq \frac{a_K}{h} (1 + o(1)),$$

$$a_K = \|F^{-2}\|^{1/2}, \quad (15)$$

$o(1) \rightarrow 0$ при $h \rightarrow \infty$.

Доказательство. Теорему можно доказать, используя утверждение 1, результат для среднеквадратичного отклонения последовательной оценки из [7], теорему 2 и неравенство (14)

ЧИСЛЕННЫЙ ПРИМЕР

В этом разделе мы приведем результаты численного эксперимента сравнения гарантированной оценки с оценками наименьших квадратов фиксированного объема линейных параметров процесса авторегрессии с условной гетероскедастичностью. Процессом для моделирования послужил двухпараметрический устойчивый ARCH-процесс:

$$x_n = \theta_1 x_{n-1} + \theta_2 x_{n-2} + \sqrt{1 + \sigma_1^2 x_{n-1}^2 + \sigma_2^2 x_{n-2}^2} \varepsilon_n, \quad (16)$$

где ε_n – н.о.р. стандартные гауссовские, $x_0 = x_{-1} = 0$. Область устойчивости $\Lambda_{(0,4;0,3)}$ этого процесса при $\sigma_1 = 0,4$, $\sigma_2 = 0,3$ приведена на рис. 1.

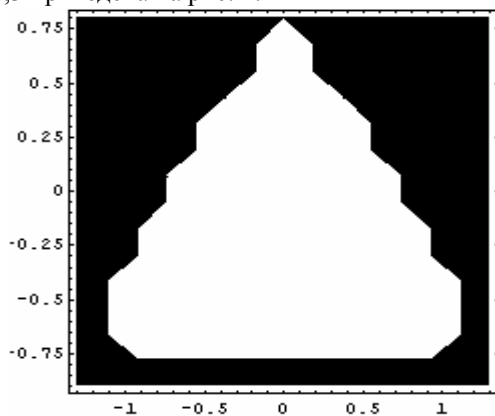


Рис. 1

Утверждение 3. Матрица $EX_k X_k'$ приближается к F по норме с геометрической скоростью при $k \rightarrow \infty$.

Доказательство. Результат основан на неравенствах (13), $\|EM\|^2 \leq E\|M\|^2$, (где M – симметричная матрица) и свойствах статистики $X_k = L_{k-1}^{-1} (x_k, \dots, x_{k-q+1})'$.

Учитывая последнее утверждение, за приближенное значение F возьмем $EX_k X_k'$. Величина $\|F^{-2}\|^{1/2}$ определяет предельное значение среднего времени остановки последовательной процедуры $E_{\theta} \tau(h)$, деленное на порог h , при $h \rightarrow \infty$ и верхнюю границу стандартного отклонения последовательной оценки $\theta^*(h)$.

Эксперимент включал 50 повторений последовательной процедуры (11), (12) для каждого порога h при различных значениях параметров θ_1, θ_2 в (16). Результаты приведены в табл. 1 и 2.

Т а б л и ц а 1

Наблюдаемые объемы выборок
в последовательном случае

θ_1	0,2	0,2	-0,5	0	0,5
θ_2	-0,5	-0,5	0	0	0
σ_1	0,4	0,4	0,1	0,1	0,1
σ_2	0,3	0,3	0,5	0,5	0,5
h	50	500	50	50	50
$\bar{\tau}(h)$	196	1902	247	211	247
$\bar{\tau}(h)/h$	3,92	3,80	4,94	4,21	4,94
$\ F^{-2}\ ^{1/2}$	3,99	3,99	4,95	4,10	4,95

Строки табл. 1 с заголовками $\bar{\tau}(h)$, $\bar{\tau}(h)/h$ и $\|F^{-2}\|^{1/2}$ обозначают соответственно наблюдаемое среднее продолжительности процедуры $\tau(h)$, это число, деленное на h , и предел $E_{\theta} \tau(h)/h$ при $h \rightarrow \infty$. Значения в табл. 1 показывают, что $\bar{\tau}(h)/h$ обеспечивает приемлемое приближение к предельному значению $\|F^{-2}\|^{1/2}$ для средних объемов выборок.

Т а б л и ц а 2

Наблюдаемые стандартные отклонения МНК
в случае фиксированной и последовательной выборок

θ_1	-0,4	-0,2	0	0,4
θ_2	0	0	0	-0,5
σ_1	0,4	0,4	0,4	0,4
σ_2	0,3	0,3	0,3	0,3
h	50	50	50	50
$SD(\theta_1^*)$	0,072	0,077	0,079	0,070
$SD(\theta_2^*)$	0,067	0,065	0,090	0,069
$hSD(\theta^*)$	4,910	5,056	5,992	4,899
$SD(\operatorname{es}\theta_1)$	0,061	0,076	0,081	0,068
$SD(\operatorname{es}\theta_2)$	0,55	0,064	0,082	0,065
$\ F^{-2}\ ^{1/2}$	4,63	4,25	4,12	4,19

Строки табл. 2 с заголовками $SD(\theta_1^*)$, $SD(\theta_2^*)$ и $hSD(\theta^*)$, обозначают соответственно наблюдаемые стандартные отклонения последовательных оценок $\theta_1^*(h)$, $\theta_2^*(h)$ и то же значение при $\theta^*(h) = (\theta_1^*(h), \theta_2^*(h))'$, умноженное на h . Строка табл. 2 с заголовком $\|F^{-2}\|^{1/2}$ обозначает теоретические верхние границы для $hE_\theta\|\theta^*(h) - \theta\|^2$, опреде-

ленные теоремой 2. Строки табл. 2 с заголовками $SD(es\theta_1)$ и $SD(es\theta_2)$ обозначают наблюдаемые стандартные отклонения обычных оценок МНК, построенные по 50 повторениям эксперимента с фиксированным объемом выборки, который был равен наблюдаемому среднему продолжительности последовательной процедуры (11), (12) при тех же значениях параметров θ_1, θ_2 в модели (16).

ЛИТЕРАТУРА

1. Андерсон Т. Статистический анализ временных рядов. М.: Мир, 1976.
2. Борисов В.З., Конев В.В. О последовательном оценивании параметров дискретных процессов // АиТ. 1977. № 10. С. 58–64.
3. Саулис Л., Статулявичус В. Предельные теоремы о больших отклонениях. Вильнюс, 1989.
4. Davydov Ju.A. The invariance principle for stationary processes // Theory probab. appl. 1970. Vol. 15. P. 487–498.
5. Feigin P.D. and Tweedie R.D. Random coefficient autoregressive process: a Markov chain analysis of stationarity and finiteness of moments// Journal of time series analysis. 1985. Vol. 6. P. 1–14.
6. Kluppelberg C. and Pergamenschikov S. The tail of the stationary autoregressive process with ARCH errors. Preprint. Munich.: Munich University of technology, Toms.: Toms State University.
7. Konev V.V. and Lai T.L. Estimators with prescribed precision in stochastic regression models // Sequential analysis. 1995. V. 14. P. 179–192.
8. Meyn S.P. and Tweedie R.D. Markov Chains and stochastic stability. London: Springer Verlag, 1996.

Статья представлена кафедрой высшей математики и математического моделирования факультета прикладной математики и кибернетики Томского государственного университета, поступила в научную редакцию «Кибернетика» 20 апреля 2004 г.